DOCUMENT RESUME

ED 433 787 HE 032 337

AUTHOR Antley, Huey W.
TITLE The Development of a Predictive Model for One-Year Freshman Retention Rate: A Macro-Approach. AIR 1999 Annual Forum Paper.
PUB DATE 1999-05-31
NOTE 28p.; Paper presented at the Annual Forum of the Association for Institutional Research (39th, Seattle, WA, May 30-June 3, 1999).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Persistence; College Entrance Examinations; *College Freshmen; Dropout Research; Higher Education; Minority Groups; Models; Nontraditional Students; On Campus Students; Part Time Students; *Predictor Variables; Private Colleges; Public Colleges; *School Holding Power; Self Evaluation (Groups); Student Attrition; *Student Characteristics
IDENTIFIERS *AIR Forum

ABSTRACT

This study investigated whether the one-year retention rate for the cohort of full-time, baccalaureate-degree-seeking, first-time freshmen could be predicted from institutional and aggregate cohort characteristics, including institutional type and control, institutional size, cohort size, average entrance exam score, percentage of part-time undergraduates, percentage of cohort residing on campus, percentage of nontraditional students, and percentage of minority students. Eight different models, based on institution type, size, and geographic location, were applied to 1995 data from 230 public and private institutions. While a completely satisfactory general model did not emerge, four of the models, all based on public four-year institutions, were particularly strong predictors of student retention. Average entrance exam score, percentage of cohort residing on campus, and percentage of nontraditional students were the most consistently significant predictor variables across the eight models. Private institutions and Research I institutions produced higher retention rates than did public institutions and other Carnegie-rated institutions. Results suggest that it is not possible to evaluate institutional effectiveness in student retention without factoring in these variables. By using these models, institutions can calculate their expected retention rate and measure actual over expected retention. (Contains 55 references.) (DB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 433 787

The Development of a Predictive Model for One-Year Freshman Retention Rate: A Macro-Approach

presented to

The Association for Institutional Research

39th Annual Forum

Seattle, Washington

May 31, 1999

Huey W. Antley, Ed.D.

Director of Institutional Research

University of Arkansas at Little Rock

2801 South University

Little Rock, Arkansas 72204

(501) 569-3302

hwantley@ualr.edu

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
D. Vura
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

HE 032 337

AIR

for Management Research, Policy Analysis, and Planning

This paper was presented at the Thirty-Ninth Annual Forum of the Association for Institutional Research held in Seattle, Washington, May 30-June 3, 1999.
This paper was reviewed by the AIR Forum Publications Committee and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC Collection of AIR Forum Papers.

Dolores Vura
Editor
Air Forum Publications

The Development of a Predictive Model for

One-Year Freshman Retention Rate: A Macro-Approach

ABSTRACT

This study investigated whether the one-year retention rate for the cohort of full-time, baccalaureate degree-seeking, first-time freshmen could be predicted from institutional and aggregate cohort characteristics, including institutional type and control, institutional size, cohort size, average entrance exam score, percentage of part-time undergraduates, percentage of cohort residing on campus, percentage of nontraditional students in the cohort, and percentage of minority students in the cohort. The study concludes that, yes, this retention rate can be predicted from these aggregate measures.

Since retention rates are influenced to a large degree by cohort characteristics, it is not possible to evaluate institutional effectiveness in student retention without factoring in these variables. Raw retention rates tell us nothing about institutional effectiveness and using these raw rates for institutional comparisons is counterproductive. The public does need to be informed regarding institutional effectiveness, and government has a right to expect accountability. However, these noble aims must be achieved through meaningful, not bogus, measures. By using the models presented in this study, many institutions can calculate their expected retention rate and develop a meaningful measure of actual over expected retention. Directions are given for calculating and evaluating this new performance indicator, the Retention Performance Ratio (RPR).

The Development of a Predictive Model for

One-Year Freshman Retention Rate: A Macro-Approach *

Over the past two decades, colleges and universities throughout the United States have focused a great deal of attention on the retention of students (Chaney & Farris, 1991; Cowart, 1987; Tinto, 1987). Initially this focus on retention was largely in response to expected and actual declines in the population of 18-year-old high school graduates. Later, many institutions began to face pressure from state legislatures anxious to implement systems of accountability, including funding tied to performance on measures such as the rate of student retention and graduation (Cave, Hanney, & Kogan, 1991; Gaither, Nedwek, & Neal, 1994; Krech, 1994). These decades of retention concern may have led to a more or less permanent change in philosophy for many institutions, resulting in institutional missions that emphasize student retention as a significant and continuing institutional goal.

Recently, the U. S. government acted to codify these retention concerns, presumably for the benefit of the general public, with the passage of the Student Right To Know and Campus Security Act (1991). In an effort to inform consumers of higher education, this act requires the collection and publication of retention rates for all institutions that receive federal funds.

However, concerns have been raised about the way in which rates of retention are evaluated and compared in the media, and about the judgements that are made, explicitly and implicitly, about institutional effectiveness based on this measure (Astin, 1993b; Astin, Tsui, and Avalos, 1996; Dey, 1990; Gaither, 1992; Krech, 1994; Mortenson, 1995, 1996a, 1996b,

---

* An earlier draft of this paper was presented at the 1st Annual Conference of the Consortium for Assessment and Planning Support (CAPS), April 18-21, 1999 in Monroe, Louisiana.

1996c; Pascarella & Terenzini, 1991). Astin (1993b) argued that "a simple retention 'rate' tells us more about who an institution admits than about how effective its retention practices are" (p. A48). Should institutions simply be more restrictive as to whom they admit in order to obtain a higher retention rate? Later, Astin, et. al (1996) expressed a clear concern that this is exactly where current practice may be leading.

> Efforts at the state level to make institutions more accountable by comparing their raw retention rates are misguided, at best, and perhaps even detrimental to state interest. The danger in such state policies is that they discourage institutions from enrolling relatively poorly prepared students in order that they be able to maximize their raw retention rates. In any state that strives to promote the quality of economic and social life for its citizens, being able to effectively educate the less well-prepared student should be given high priority, since such students pose the greatest risk of eventually becoming dependent on the state.... Therefore, any state policy that discourages institutions from admitting and educating underprepared students basically works in opposition to long-term best state interests [emphasis added]. (p. 16)

Given a nationwide effort to improve retention, why are national retention rates not rising? In fact, some national and regional studies report that, overall, retention has remained relatively unchanged over many years (Gaither, 1992; Seidman, 1996; Tinto, 1982b). The American College Testing Program (ACT) has reported that, in the 13 years from 1983 to 1996, the overall rate of retention actually decreased by 2.4 percentage points (Geraghty, 1996; Mortenson, 1996a).

Regarding the implementation of performance indicators, such as the institutional retention rate, Cave et al. (1991) warned that "the partial nature of most indicators leaves

ample scope for strategic behavior on the part of the unit being appraised. In other words, institutions can manipulate inputs to achieve outputs without improving processes" (p. 34). Could it be that we are simply shifting high-risk students from one type of institution to another so that some can appear to have become more productive?

If an institution wants to know how well it is doing with the students it already serves, what should it do? Unfortunately, despite the vast body of retention research, there is little guidance for practitioners attempting to set appropriate institutional goals for student retention, although studies have noted differences in retention rates among different types of institutions (Astin, 1993b; Mortenson, 1996a; Smith, 1997). While these studies have used descriptive statistics to describe differences in retention rates and student characteristics, they have not made use of inferential statistics to make predictions from these aggregate institutional data.

Research Literature

Most of the published research deals with the individual student—a micro view. This is understandable since a primary objective in studying the attrition phenomenon has been to devise ways to avoid this outcome. This micro-level research has been fruitful and a number of micro-level prediction models have been developed (Astin et al., 1996; Bean, 1980, 1986; Cabrera, Castañeda, Nora, & Hengstler, 1992; Dey, 1990; Pascarella & Terenzini, 1980). However, the macro-view of student retention has received little research focus until recently (Astin, 1993b; Gaither, 1992; Mortenson, 1995).

The problem of student attrition in higher education has been researched for many years, with studies dating back to the 1920s (Gavit, 1925; Minnesota University, 1924). While this attrition research has yielded a host of potentially relevant variables for

understanding the dynamics of individual enrollment and withdrawal decisions (Astin, 1993c; Bean, 1982; Pascarella & Terenzini, 1991; Tinto, 1975, 1982a, 1982b), the research has not provided a simple model for predicting institutional retention rates from aggregate data. If an expected rate of retention can be derived, then institutions can be judged by how well they perform against a reasonable expectation, not by how they perform against other schools. This study sought to establish a useful model for this purpose based on readily available aggregate data.

Early persistence research tended to be descriptive (Pascarella & Terenzini, 1991); later, a number of theorists posited models of persistence behavior that then formed the basis of subsequent research (Bean, 1980, 1982; Cabrera et al., 1992; Pascarella & Terenzini, 1980; Spady, 1970, 1971; Tinto, 1975, 1985). Quite rightly, these early studies focused on individual persistence decisions and are appropriately classified as micro-analyses. These micro-studies helped to identify individual students at risk of attrition and helped program developers to fashion interventions intended to avert these undesirable outcomes (Cowart, 1987).

However, with the recent emphasis on institutional assessment and accountability (Cave et al., 1991; Gaither et al., 1994; Krech, 1994), as institutional retention rates have begun to be widely published and compared, the focus of concern has shifted to the institution level. Recently, a few macro-level studies have appeared that view persistence decisions in aggregate (Astin, 1993b; Astin et al., 1996; Dey, 1990; Mortenson, 1995, 1996a, 1996b, 1996c; Smith, 1997). Astin et al. (1996) developed a predictive model for institutional retention rate, but used micro-level data collected from a sample of students. He suggested

that, to obtain an expected institutional retention rate, the model coefficients be applied to individual student data, followed by the averaging of the individual probabilities.

Based on this body of literature, key variables that have proven useful in predicting micro-level retention were evaluated with regard to their potential usefulness as aggregate measures in a macro-analysis. The conclusion from this review was that the following variables would prove useful in this study: institutional type and control, institutional size, cohort size, average entrance exam score, percentage of part-time undergraduates, percentage of cohort residing on campus, percentage of nontraditional students in the cohort, and percentage of minority students in the cohort.

Research Question

This study sought to develop a predictive model that would identify an expected rate of retention given a particular entering freshman cohort. With such a prediction, a school can be evaluated based on how well its actual retention stacks up against its expected retention. In particular, this study explored whether the one-year retention rate for the cohort of full-time, baccalaureate degree-seeking, first-time freshmen could be predicted from institutional and aggregate cohort characteristics, including institutional type and control, institutional size, cohort size, average entrance exam score, percentage of part-time undergraduates, percentage of cohort residing on campus, percentage of nontraditional students in the cohort, and percentage of minority students in the cohort.

Data Collection and Analysis

This study made use of a dataset collected by the Consortium for Student Retention Data Exchange (CSRDE) from 230 public and private institutions in the U.S. regarding their 1995 entering freshman cohorts. These data were recoded where necessary to fit the

requirements of a multiple regression analysis for the purpose of developing coefficients for the prediction of an expected retention rate. Statistical significance at all stages of the analysis was based on an alpha-level of $\alpha = 0.05$ in a two-tailed test.

This study extended the movement toward aggregate measures in retention research by moving macro-analysis from the descriptive to the inferential realm. Following the guide of macroeconomics (Boulding, 1966; Minford, 1992; Shapiro, 1970), this study sought to look at groups of individuals and to predict group behavior, recognizing that while individual behavior can rarely be predicted, group behavior tends to be quite predictable (Minford, 1992). This macro-analysis approach has yielded coefficients that can be applied directly to aggregate measures to predict an expected retention rate.

Dichotomous variables. All of the dichotomous variables presented in Table 1 were coded one (1) or zero (0): one for yes, meaning the institution is considered a member of the group identified by the variable; and zero for no, meaning the institution is not considered a member of the group identified by the variable.

Table 1. Dichotomous Factors and Variables

Factor	Variable Label	Number	Percent of Total
Control	Public	158	70.2
	Private	42	18.7
	Other	25	11.1
Type	Baccalaureate	23	10.2
	Masters	88	39.1
	Doctoral	36	16.0
	Research 2	17	7.6
	Research 1	36	16.0
	Other	25	11.1
Size	Small	60	26.7
	Medium	97	43.1
	Large	68	30.2
Region (publics only)	Heartland	53	33.5
	Northeast	10	6.3
	South	52	32.9
	West	43	27.2

Non-Dichotomous Variables. Most of the non-dichotomous variables represented percentages, all of which were percentages of the entering freshman cohort, with the exception of Percentage Part-Time Undergraduates. For each of the freshman-based percentages, a new variable was constructed by multiplying Cohort Size by the percentage divided by 100. The name for each new variable was constructed by adding the prefix "N-" to the percentage variable name, identifying the variable as the number equivalent of the percentage. This number equivalent was used in component-based model testing as described below.

Table 2. Non-Dichotomous Factors and Variables

Factor	Variable Label	Mean	SD	Min	Max
Cohort Size	Count	1630	1436	53	10947
Average Entrance Exam Score	ACT	21.3	2.4	15.0	27.6
	SAT	964	103	708	1217
	ACT/SAT	21.1	2.4	15.0	27.6
Percentage Part-Time Undergraduates	Part-time	21	15	1	84
Percentage Minority	Minority	17	16	0	99
	N-Minority	238	256	0	1372
Percentage Nontraditional	Non-Trad	2	4	0	28
	N-Non-Trad	23	38	0	246
Percentage Living On Campus	On Campus	73	23	0	100
	N-On-Campus	1219	1117	0	5535
Retention Rate	Retention	75	9	48	97
	N-Retained	1277	1231	31	8976

Institutional type and control. Astin (1993c) concluded that the effects of institutional type are indirect. More direct influences include the types of student-student and student-faculty interactions that occur in, and perhaps are fostered by, these differing environments. However, since in this study these more direct variables were unavailable, the institutional type served as a proxy for these effects. The CSRDE data categorized participating institutions by the Carnegie classifications: Research 1, Research 2, Doctoral, Masters,

Baccalaureate, and included both public and private institutions ranging in size from under 5,000 to over 18,000 students. While these classifications may not directly influence the specific dynamics of student retention, it could be argued that they do represent some relevant differences in size, mission, and focus on undergraduates, particularly freshmen, which may impact both the type of cohort selected and the kind of collegiate experience the cohort encounters. In addition, two dichotomous variables were constructed to designate institutional control: "public" and "private."

Institutional size. Institutional size has been shown in previous studies to be associated with retention. Astin (1993c), replicating his previous studies (Astin, 1975, 1982), reported that institutional size is negatively correlated with retention. Institutional size was identified in the CSRDE report by assignment to one of three categories: small, medium, or large. The definition of the categorization was as follows: small – less than 5,000 students; medium – 5,000 to 17,999 students; and large – 18,000 or more students.

Region. A close examination of the CSRDE dataset revealed that representation was not equal across all parts of the United States. This finding led to the decision to examine separate models by region. To facilitate this analysis, an attempt was made to identify the location of each institution by matching the institution name in the CSRDE dataset to the U.S. Department of Education - Integrated Postsecondary Education Data System (IPEDS) for Fall 1995. This match was accomplished successfully for all of the public institutions; however, the frequent duplication of names among private institutions and the lack of location data in the CSRDE dataset prevented the determination of consistent matches to IPEDS for the private institutions.

Cohort size. Cohort size was used as both a control variable and a potential factor in retention. Since the size of the freshman cohort is somewhat indicative of the overall size of the institution, this variable, labeled "count," was also considered to have value as a proxy for institutional size. The CSRDE survey instructions, referring to IPEDS instructions, defined the first-time freshman cohort to be inclusive of all first-time, full-time, baccalaureate degree-seeking freshmen. This was inclusive of students who earned credits while still in high school and students who enrolled for the first time in the prior summer term.

Average entrance exam score. Numerous studies (Astin, 1971, 1972, 1975, 1977, 1993b, 1993c; Astin et al., 1996; Bean, 1980, 1982, 1986; Dey, 1990; Gaither, 1992; Gillespie & Noble, 1992; House, 1994; Mortenson, 1995, 1996a, 1996b, 1996c; Pascarella & Terenzini, 1980, 1991; Tinto, 1975) identify a strong association between academic preparation and retention. In these studies, high school GPA or high school rank is generally identified as the best single predictor, but scores on entrance exams (such as the ACT and SAT) are also positively associated with retention (Astin, 1993c; House, 1994). The CSRDE dataset did not contain an average high school GPA for the cohorts; therefore, in this study, the average entrance exam score was used as the sole measure of academic preparation.

Some CSRDE institutions reported average ACT score and some average SAT score, depending on which test was the primary score used in the admissions process. Other schools reported both scores but did not provide data regarding the number of students taking each test. A new variable, ACT/SAT, was constructed for entrance exam score and populated with either the ACT score, an ACT equivalent for the SAT score, or a simple average of the ACT score and the ACT-equivalent SAT score. Where both scores were reported, the following procedure was used to set the value of ACT/SAT.

In consultation with the American College Testing Service (ACT), a table was constructed identifying the status of each state as regards these two tests. For states that primarily use ACT, the ACT score was used. For states that primarily use SAT, the ACT-equivalent SAT score was used. For states that commonly use both scores, the simple average of the ACT and ACT-equivalent SAT was used.

Percentage part-time undergraduates. Part-time students have been identified as an at-risk group for attrition (Astin, 1993c; Pascarella & Terenzini, 1991). Although the cohort under investigation was the group of full-time, first-time freshmen, an academic environment dominated by part-time students may limit opportunities for social and academic integration.

Percentage minority. Minority students in majority white institutions have been shown to be at higher risk of attrition (Astin, 1993c; Astin et al., 1996; Pascarella & Terenzini, 1991). This risk may be the result of difficulties faced in integrating socially and academically with the majority white population of students and faculty. Also, a higher percentage of these students may be first-generation college students, another risk factor. The CSRDE survey asked for the percentage of the entering freshman cohort that consisted of underrepresented minorities.

Percentage nontraditional. Nontraditional students have been shown in previous studies to have lower retention rates than do traditional-college-age students (Astin, 1993c; Pascarella & Terenzini, 1991). The CSRDE survey asked for the percentage of the entering freshman cohort that was 24 years or older in age.

Percentage on-campus. In numerous studies (Astin, 1975, 1977, 1982, 1993c; Pascarella & Terenzini, 1991), residence on campus has been shown to improve the chance of retention for an individual student. Consequently, the aggregate percentage of entering

freshmen living on campus should be positively associated with the one-year retention rate for the matriculating freshmen cohort.

Method

The following method was employed in the development of each of the following eight models: Model 1 - a general model; Model 2 - public institutions; Model 3 - private institutions; Model 4 - public institutions in the heartland states; Model 5 - public institutions in the northeastern states; Model 6 - public institutions in the southern states; Model 7 - public institutions in the western states; and Model 8 - public institutions in the heartland and southern states.

First, the independent relationship of each predictor variable with the dependent variable, retention rate, was examined. Variables with promising independent relationships were included in a series of trial multiple regressions. In subsequent iterations of the multiple regression analysis, variables with weak significance were removed until all remaining variables, and the model as a whole, reached the specified significance level, $p < 0.05$.

As a final step, the model was modified to address concerns about regression analysis using ratio variables with common denominators (Atchley, Gaskins, & Anderson, 1976; Bollen & Ward, 1980; Chayes, 1971; Kuh & Meyer, 1955; Pearson, 1897; Pendleton, Warren, & Chang, 1980; Schuessler, 1974; Uslaner, 1976, 1977; Yule, 1910). Several of the percentages used in this study, including the dependent variable and several of the independent variables, were all based on a common denominator. To deal with these concerns, the model was tested using component variables substituted for the percentage variables, as recommended by Freeman and Kronenfeld (1973), with the cohort size entered as a control variable and the number retained specified as the dependent variable. When

statistical significance was found for both the percentage-based model and the component-based model, the model was considered sound.

Findings

While all of the resulting models were statistically significant ($p < 0.001$), four could be described as relatively strong and four as relatively weak.

Stronger models. The stronger models were all based on public four-year institutions: Model 2 - all public institutions; Model 4 - heartland states public institutions; Model 6 - southern states public institutions; Model 8 - combined heartland and southern states institutions. The CSRDE dataset contained 25% of the public four-year institutions contained in the comparable IPEDS dataset. These models appear to be relatively strong because they are based on apparently good samples of their respective populations, they coincide with earlier research findings, and they are relatively consistent with one another.

Table 3. Comparison of Coefficients and Model Statistics for Stronger Models: 2, 4, 6 and 8

Independent Variable	Model 2 Public	Model 4 Heartland	Model 6 South	Model 8 Heartland & South
Intercept	43.358 ***	38.927 ***	35.067 **	39.401 ***
Research 1	+7.077 ***		+10.37 ***	+8.272 ***
Count		+0.001 **		
ACT/SAT	+1.270 ***	+1.626 ***	+1.156 *	+1.378 ***
Minority		-0.325 *		
Non-Trad	-0.773 **			-0.624 *
On Campus	+0.068 *		+0.200 ***	+0.085 **
Model Prob	0.000	0.000	0.000	0.000
R-Squared	0.575	0.571	0.722	0.618

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

As can be seen in Table 3, there are some distinct similarities among the four models. First, they have similar intercepts, ranging only 8 points from 35 to 43. Then, all the models but Model 4 show a 7 to 10 point increase in retention for the Research 1 institutions. Model

4, by contrast, increases retention based on cohort size; since Research 1 institutions are presumably the largest of the institutional types, this variable "count" may serve as a proxy for "Research 1" in Model 4. Next, all four models have a similar coefficient for ACT/SAT, ranging from 1.2 to 1.6. Clearly, each point of ACT/SAT yields a point or more of retention improvement. Further, all the models but Model 4 add points for the percentage on campus, with coefficients ranging from 0.068 to 0.200. This means that each 10-point gain in percentage living on campus yields almost a point or more gain in retention. The R-Squareds for these models suggest that they all explain more than half the variation in retention rate with Model 6 for the South explaining almost three-quarters of the variation.

Weaker models. The other four models are relatively weak: Model 1 - the general model; Model 3 - all private institutions; Model 5 - northeastern states public institutions; and Model 7 - western states public institutions.

Table 4. Comparison of Coefficients and Model Statistics for Weaker Models: 1, 3, 5, and 7

Independent Variable	Model 1 General	Model 3 Private	Model 5 Northeast	Model 7 West
Intercept	40.875 ***	77.843 ***	31.315 *	77.950 ***
Public	-4.388 **			
Doctoral				-8.188 ***
Small				-15.532 ***
Research 1	+7.629 ***			
Count		+0.006 *		
ACT/SAT	+1.503 ***		2.292 ***	
Minority		-0.222 **		
Non-Trad				-1.090 **
On Campus	+0.070 **			
Model Prob	0.000	0.000	0.001	0.000
R-Squared	0.438	0.348	0.774	0.709

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

These models are weak because they are based on suspect samples, they are not consistent with earlier research findings, and they are not consistent with one another. The CSRDE dataset contained only 2% of the private four-year institutions contained in the comparable IPEDS dataset and only 8% of the northeastern public four-year institutions. Regarding Model 7 for western states public institutions, although 39% of the western public four-year institutions supplied data to CSRDE, most of these cases were missing key variables.

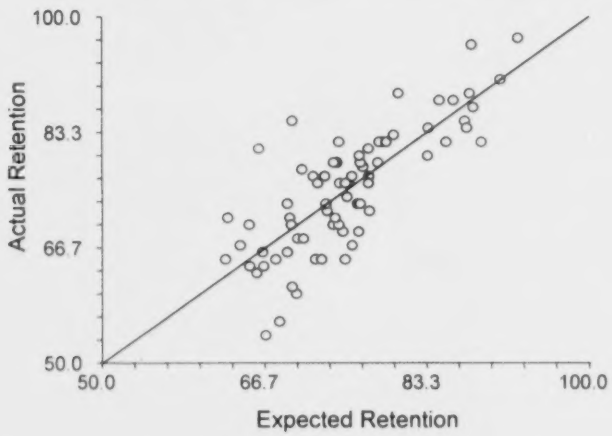
A comparison of the coefficients for Models 1, 3, 5, and 7 displayed in Table 4 reveals no consistency at all among the models. First, the intercepts range from 78 to 30, a range of 48 points. Next, there is no single variable that all the models share in common. Then, only ACT/SAT appears in more than one of the four models. The general model is most like the stronger models 2, 4, 6, and 8 presented earlier; however, it suffers in that it does not contain a good sample of the private institutions. Next, the model that emerged for the private institutions is not satisfying theoretically, since it uses only two variables, count and percentage minority, and yields the lowest R-Squared of all the models. Finally, Model 5 for the northeast states and Model 7 for the western states both suffer from poor samples.

Conclusions

The purpose of this study was to explore whether the one-year retention rate for the cohort of full-time, baccalaureate degree-seeking, first-time freshmen could be predicted from institutional and aggregate cohort characteristics. The conclusion drawn from the findings presented above is that, yes, the one-year retention rate can be predicted from these aggregate measures.

Effective models. While a completely satisfactory general model did not emerge from this study, very good models emerged for several sub-sets of the population of interest, and there is good reason to believe that, with the availability of a better sample, a completely satisfactory general model can be developed as well. The strongest model was Model 8 for public institutions in the heartland and southern states. Figure 1 illustrates the effectiveness of this model in predicting the one-year retention rate for these CSRDE institutions.

Figure 1. Scatter Plot of Predicted to Actual Retention Rate for Model 8



Predictor variables. As expected, student pre-college characteristics account for much of the variation in retention rate. Average entrance exam score, percentage of cohort residing on campus, and percentage of nontraditional students in the cohort were the most consistently significant predictor variables across the eight models. Beyond the influence of these variables, the study's findings also support earlier findings suggesting that public and private institutions, when given apparently equivalent cohorts, produce differing results in retention (Astin, 1993c; Astin et al., 1996; Pascarella & Terenzini, 1991). In short, private institutions produce higher retention rates. However, this study cannot conclude that this difference is based solely on the nature of the private higher education experience.

Further, this study strongly suggests that Research 1 institutions are unique with regard to retention. However, once again, this study cannot conclude whether this superior retention performance is the result of a unique collegiate experience or the result of differential cohorts possessing characteristics as yet unmeasured.

Implications

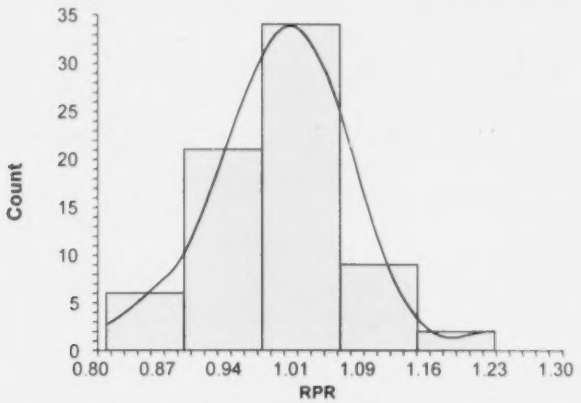
Since retention rates are influenced to a large degree by cohort characteristics, it is not possible to evaluate institutional effectiveness in student retention without factoring in these variables. As Astin (1993a) so aptly put it, "Any educational assessment project is incomplete unless it includes data on student inputs" (p. 18). The public does need to be informed regarding institutional effectiveness, and government has a right to expect accountability. However, these noble aims must be achieved through meaningful, not bogus, measures. Further, at the institutional level, administrators cannot assess the value of resources spent in retention efforts without first understanding the natural implications of their admissions policies. By using the models presented in this study, many institutions can calculate their expected retention rate and develop a meaningful measure of actual over expected retention.

Application

To apply the results of this study, researchers and administrators interested in evaluating retention rates must calculate the "retention performance ratio" (RPR) for each school of interest. The RPR is calculated as follows: (1) select the most appropriate model; (2) obtain the aggregate measures used in the model; (3) calculate the expected retention rate using the constant and model coefficients; and (4) divide the actual retention rate by the expected retention rate to obtain the "retention performance ratio." Schools that are performing at expectation will have a retention performance ratio (RPR) of about 1.0. Figure

2, which displays the distribution of the RPRs for the CSRDE public universities in the heartland and southern states, reveals that, as expected, the mean for the distribution is 1.0.

Figure 2. Histogram of Retention Performance Ratio for Model 8 - Heartland and Southern States



The standard deviation of the distribution of RPR could serve as a convenient demarcation of retention performance. For example, schools performing from one standard deviation below to one standard deviation above the mean could be considered as "performing at the expected level." Institutions with an RPR greater than one standard deviation above the mean but less than two standard deviations above the mean could be considered as "performing slightly above expectations." Further, institutions with an RPR greater than two standard deviations above the mean could be considered as "clearly performing above expectations." Similarly, these boundaries could define schools "performing slightly below expectations" and those "clearly performing below expectations."

Table 5. Retention Performance Ratio (RPR) Performance Categorization for Model 8

Performance Category	RPR Range	Count	Percent
Clearly Performing Above Expectations	Greater than 1.154	2	3%
Performing Slightly Above Expectations	1.078 to 1.154	5	7%
Performing at the Expected Level	0.923 to 1.077	56	78%
Performing Slightly Below Expectations	0.846 to 0.922	6	8%
Clearly Performing Below Expectations	Less than 0.846	3	4%

When this rubric is applied to Model 8 for the heartland and southern states, the model's standard deviation of 0.077 results in the demarcation of performance categories as presented in Table 5. The standard deviation of RPR for the four strong models ranges only from 0.068 to 0.077, averaging 0.075.

Future Research

Additional research is needed in two major areas to extend and confirm the findings of this study. First, the study should be replicated with a new and more complete dataset. This opportunity should be forthcoming in the near future as IPEDS implements a full-scale collection of persistence and graduation data. Second, model coefficients and RPR performance boundaries should be used to evaluate schools to identify those that are clearly performing above and below expectations. A detailed analysis of the methods employed in these schools may be instructive with regard to truly effective interventions, or perhaps these schools will be representative of unique environments that are impacting the retention performance for good or ill.

Further research into expected retention rates is critically important. First, the current study shows very promising results that deserve to be confirmed and extended. Second, the implications of such findings are great. Although a number of voices have argued persuasively for a revised approach to performance assessment, particularly in the area of student retention (Astin, 1993b; Astin et al., 1996; Dey, 1990; Mortenson, 1996c), the momentum behind the present system of raw retention rate comparisons is great.

This study, in concert with the work of others, moves us in the right direction. However, much work remains to be done before valid comparisons can lead to real insight into the performance of institutions with regard to student retention.

REFERENCES

Astin, A. W. (1971). Predicting academic performance in college. New York: Free Press.

Astin, A. W. (1972). College dropouts: A national profile. ACE Research Reports. Washington, DC: American Council on Education.

Astin, A. W. (1975). Preventing students from dropping out. San Francisco: Jossey-Bass.

Astin, A. W. (1977). Four critical years: Effects of college on beliefs, attitudes, and knowledge. San Francisco: Jossey-Bass.

Astin, A. W. (1982). Minorities in American higher education. San Francisco: Jossey-Bass.

Astin, A. W. (1993a). Assessment for excellence: The philosophy and practice of assessment and evaluation in higher education. Phoenix: Oryx Press.

Astin, A. W. (1993b, September 22). College retention rates are often misleading. Chronicle of Higher Education, p. A48.

Astin, A. W. (1993c). What matters in college?: Four critical years revisited. San Francisco: Jossey-Bass.

Astin, A. W., Tsui, L., & Avalos, J. (1996). Degree attainment rates at American colleges and universities: Effects of race, gender, and institutional type. Los Angeles, CA: Higher Education Research Institute. (ERIC Document Reproduction Service No. ED 400 749)

Atchley, W., Gaskins, C., & Anderson, D. (1976). Statistical properties of ratios. Systematic Zoology, 25, 137-148.

Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. Research in Higher Education, 12(2), 155-187.

Bean, J. P. (1982). Conceptual models of student attrition: How theory can help the institutional researcher. In E. Pascarella (Ed.), New directions for institutional research: Vol. 9 (4). Studying student attrition (pp. 17-33). San Francisco: Jossey-Bass.

Bean, J. P. (1986). Assessing and reducing attrition. In D. Hossler (Ed.), New directions for higher education: No. 53. Managing college enrollments (pp. 47-61). San Francisco: Jossey-Bass.

Bollen, K. A., & Ward, S. (1980). Ratio variables in aggregate data analysis: There uses, problems and alternatives. In Borgatta, E. F. & Jackson, D. J. (Eds.), Aggregate data (pp. 60-79). Beverly Hills, CA: Sage.

Boulding, K. E. (1966). Economic analysis (4th ed., Vol. 2). New York: Harper and Row.

Cabrera, A. F., Castañeda, M. B., Nora, A., & Hengstler, D. (1992). The convergence between two theories of college persistence. Journal of Higher Education, 63(2), 143-164.

Cave, M., Hanney, S., & Kogan, M. (1991). The use of performance indicators in higher education: A critical analysis of developing practice. London: Jessica Kingsley Publishers.

Chaney, B. & Farris, E. (1991). Survey on retention at higher education institutions (Higher Education Surveys Report, Survey No. 14). Rockville, MD: Westat, Inc.). (ERIC Document Reproduction Service No. ED 342 334)

Chayes, F. (1971). Ratio correlation. Chicago: University of Chicago Press.

Cowart, S. C. (1987). What works in student retention in state colleges and universities. Washington, DC: American College Testing Program and the American Association of State Colleges and Universities. (ERIC Document Reproduction Service No. ED 347 928)

Dey, E. L. (1990). Evaluating college student retention: Comparative national data from the 1981-1984 entering freshman classes. Paper presented at the Annual Meeting of the American Educational Research Association, Boston, MA. (ERIC Document Reproduction Service No. ED 319 320)

Freeman, J. H. & Kronenfeld, J. E., (1973). Problems of definitional dependency: The case of administrative intensity. Social Forces, 52, 108-121.

Gaither, G. H. (1992). Persistence patterns in public higher education: The case of Texas. College and University, 67(1), 245-252.

Gaither, G., Nedwek, B. P, & Neal, J. E. (1994). Measuring up: The promises and pitfalls of performance indicators in higher education (ASHE-ERIC Higher Education Rep. No. 5). Washington, DC: The George Washington University, Graduate School of Education and Human Development.

Gavit, J. P. (1925). College. New York: Harcourt, Brace and Co.

Geraghty, M. (1996, July 19). Data show more students quitting college before sophomore year. The Chronicle of Higher Education, p. A35.

Gillespie, M., & Noble, J. (1992). Factors affecting student persistence: A longitudinal study (ACT Research Report Series 92-4). Iowa City, IA: The American College Testing Program.

House, J. D. (1994, November). College grade outcomes and attrition: An exploratory study of noncognitive variables and academic background as predictors. Paper presented at the Illinois Association for Institutional Research Annual Meeting, Lake Shelbyville, IL. (ERIC Document Reproduction Service No. ED 390 319)

Krech, A. S. (1994, July). Accountability and the movement to performance indicators: The South Carolina experience and guidelines for institutional effectiveness. Materials from a presentation at the Summer Data Conference of the National Center for Education Statistics, Arlington, VA. (ERIC Document Reproduction Service No. ED 371 784)

Kuh, E., & Meyer, J. R. (1955). Correlation and regression estimates when the data are ratios. Econometrica, 23, 400-416.

Minford, P. (1992). Rational expectations macroeconomics: An introductory handbook (2nd ed.). Oxford UK and Cambridge USA: Blackwell.

Minnesota University (1924). Report of Survey Commission VI: Student mortality. Bulletin of Minnesota University, 4.

Mortenson, T. G. (1995, March). Institutional graduation rates. Postsecondary Education Opportunity, 33, 9-14.

Mortenson, T. G. (1996a, February). Freshman-to-sophomore persistence by institution level, control, and academic selectivity. Postsecondary Education Opportunity, 44, 1-9.

Mortenson, T. G. (1996b, March). Institutional graduation rates by degree level, control, and academic selectivity. Postsecondary Education Opportunity, 45, 1-9.

Mortenson, T. G. (1996c, December). Freshman-to-sophomore persistence by institution level, control, and academic selectivity. Postsecondary Education Opportunity, 54, 1-4.

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. Journal of Higher Education, 51 (1), 540-552.

Pascarella, E. T., & Terenzini, P. T. (1991). How college affects students: Findings and insights from twenty years of research. San Francisco: Jossey-Bass.

Pearson, K. (1897). On a form of spurious correlation which may arise when indices are used in the measurement of organs. Proceedings of the Royal Society of London, 60, 489-498.

Pendleton, B. F., Warren, R. D., & Chang, H. C. (1980). Correlated denominators in multiple regression and change analysis. In Borgatta, E. F. and Jackson, D. J. (Eds.), Aggregate data (pp. 80-103). Beverly Hills, CA: Sage.

Schuessler, K. (1974). Analysis of ratio variables: Opportunities and pitfalls. American Journal of Sociology, 80, 379-396.

Seidman, A. (1996). Retention revisited: R = E, Id + E & In, Iv. College and University, 71, 18-20.

Shapiro, E. (1970). Macroeconomic analysis (2nd ed., W. J. Baumel, Ed.). New York: Harcourt, Brace, & World.

Smith, T. Y. (Ed.). (1997). 1996-97 CSRDE Report: The retention and graduation rates of 1988-1995 entering freshmen cohorts in 230 U.S. colleges and universities. Norman, Oklahoma: University of Oklahoma, Office of Institutional Research.

Spady, W. (1970). Dropouts from higher education: An interdisciplinary review and synthesis. Interchange, 1, 64-85.

Spady, W. (1971). Dropouts from higher education: Toward an empirical model. Interchange, 2(3), 38-62.

Student Right To Know and Campus Security Act, 20 U.S.C. § 1092 (1991).

Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. Review of Educational Research, 45, 89-125.

Tinto, V. (1982a). Defining dropout: A matter of perspective. In E. Pascarella (Ed.), New Directions for Institutional Research: No. 36. Studying Student Attrition (pp. 3-15). San Francisco: Jossey-Bass.

Tinto, V. (1982b). Limits of theory and practice in student attrition. Journal of Higher Education, 53, 687-700.

Tinto, V. (1985). Theories of student departure revisited. In J. C. Smart (Ed.), Higher education: Vol. 2. Handbook of theory and research (pp. 359-384). New York: Agathon Press.

Tinto, V. (1987). Leaving college. Chicago and London: The University of Chicago Press.

Uslaner, E. (1976). The pitfalls of per capita. American Journal of Political Science, 20, 125-133.

Uslaner, E. (1977). Straight lines and straight thinking: Can all those econometricians be wrong? American Journal of Political Science, 21, 183-191.

Yule, G. U. (1910). On the interpretation of correlations between indices or ratios. Journal of the Royal Statistical Society, 73, 644-647.